

# Joint Genome Institute's Automation Approach and History

June 27, 2006

**Simon Roberts**

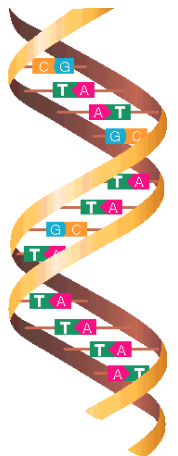
*(Production Instrumentation Supervisor)*

This work was performed under the auspices of the US Department of Energy's Office of Science, Biological and Environmental Research Program, and by the University of California, Lawrence Livermore National Laboratory under Contract No. W-7405-Eng-48, Lawrence Berkeley National Laboratory under contract No. DE-AC02-05CH11231 and Los Alamos National Laboratory under contract No. DE-AC02-06NA25396

**LBNL - 60620**

# Agenda

- Brief overview of the how the Joint Genome Institute came into existence.
- Overview of DNA sequencing production line at the JGI.
- How our throughput has increased since 1999 to become a high through-put sequencing facility.
- Some instrumentation improvement highlights along the way and how they are used.
- Review our approach to successful selection & implementation of new instruments to meet our needs.



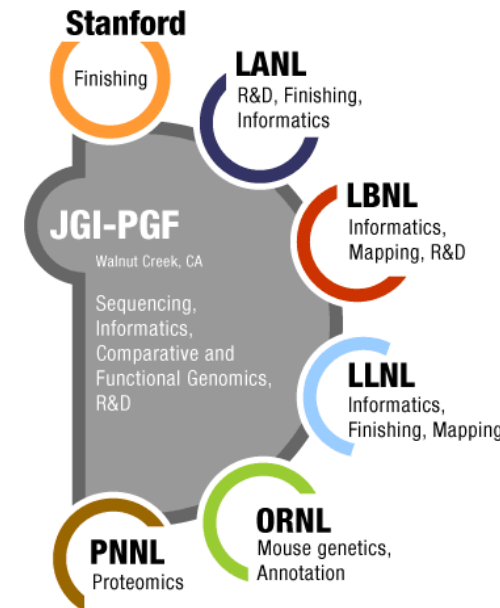


# DOE JGI

## Production Genomics Facility



**Opened in 1999**  
**~240 UC Employees**  
**60,000 sf**  
**~\$66M Annual Budget**



The DOE Joint Genome Institute (JGI) is a "virtual institute" that integrates the sequencing and analytical activities of six partner institutions:

### ***Mission:***

***DOE JGI collaborates with DOE national laboratories and community users, to advance genome science in support of the DOE missions of clean bio-energy, carbon cycling, and bioremediation.***

# Important Dates in DOE Genomics

- **1986** DOE announces Human Genome Initiative. With \$5.3 million, pilot projects begin at DOE national laboratories to develop critical resources and technologies.
- **1990** DOE & NIH present their joint HGP plan to Congress. The 15-year project formally begins.
- **1997** DOE creates the JGI uniting activities at DOE human genome centers.
- **1999** JGI opens the Production Genomics Facility (PGF) in Walnut Creek, staff from LLNL & LBNL.
- **2000** HGP leaders & President Clinton announce the completion of a “working draft..the first great technological triumph of the 21st century.”
- **2003** HGP completed and published.



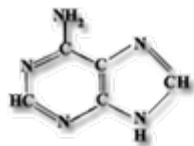
# What IS a Genome???

A **GENOME** is all of a living thing's genetic material.

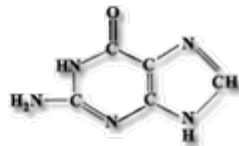
The genetic material is **DNA**  
(**D**eoxyribo**N**ucleic **A**cid)

DNA, a double helical molecule, is made up of four nucleotide "letters":

A--



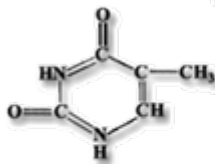
Adenine



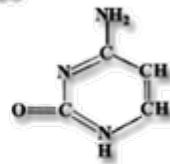
Guanine

Purines

T--



Thymine

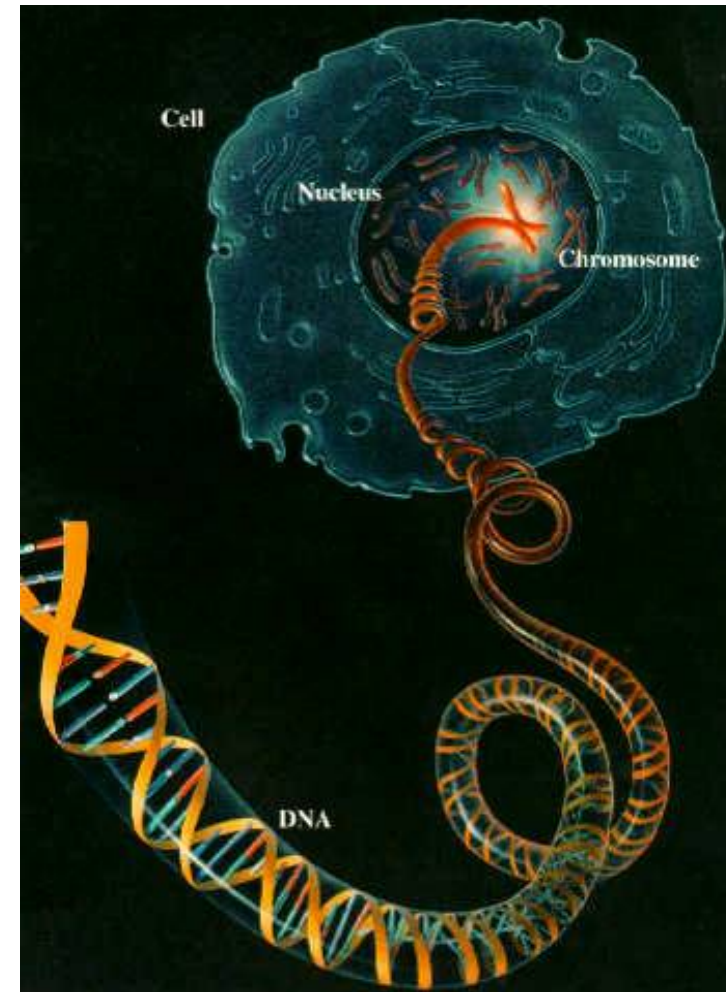


Cytosine

Pyrimidines

G--

C--

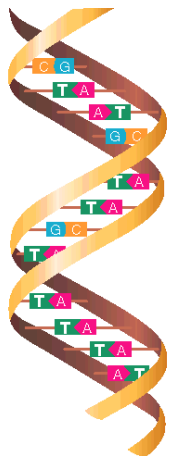


# JGI Production Mission

**“Produce high quality cost efficient  
‘assemble-able’ sequence in a safe  
environment.”**

## FY 2006 Goals

- 52 million lanes = 35 billion bases



# How Sequencing is Done

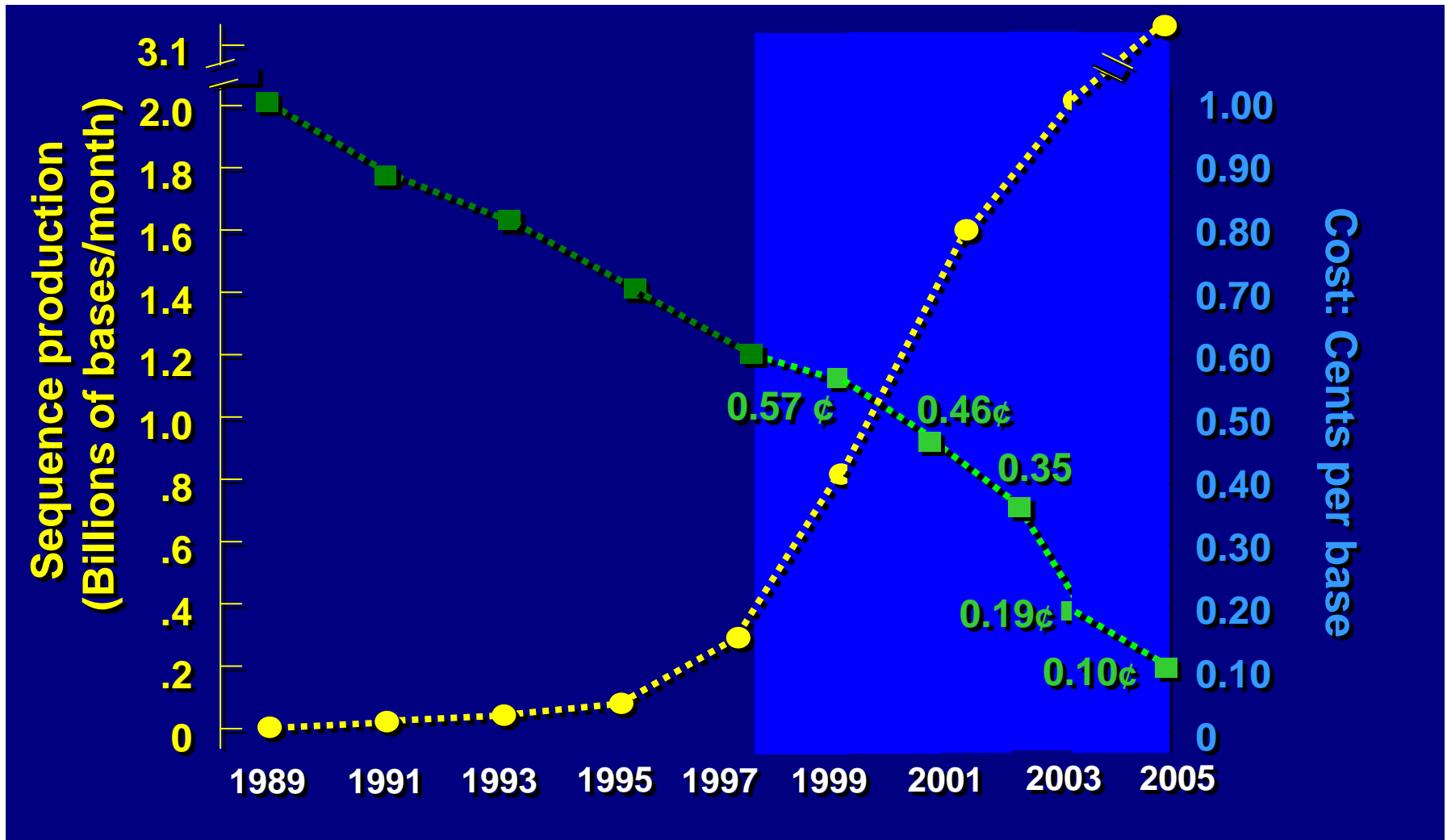
## Production Line Overview

- **Library Creation**
  - Shearing the DNA (Genomic Solutions Hydroshear)
  - Insertion of Fragments into Plasmid (Ligation)
  - Transformation (Electroporation)
  - Subcloning the Sheared Fragment (Plating)
  - **Colony Picking (Genetix QPix)**
- **Production Sequencing**
  - Lysing the Cell (Matrix PlateMate)
  - Rolling Circle Amplification (MultiDrop Micro)
  - **Sequencing Chemistry (CyBio Vario)**
  - Post Sequencing Reaction Cleanup (BioMek FX)
  - **Capillary Sequencing (ABI 3730 & MB 4500)**
- **Assembly & QA**
  - Assembly (Phrap, JAZZ)
  - Quality Assessment (Phred Q20)



# DOE JGI

## Production Sequencing Efficiencies



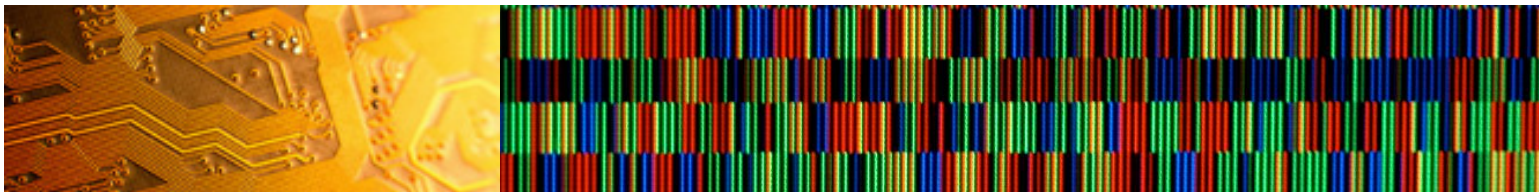
# “Moore’s Law” of DNA Sequencing

- **April 2002:** 1 Gb/month
- **January 2004:** 2 Gb/month
- **July 2004:** 2.5 Gb/month
- **March 2005:** 3.1 Gb/month

(equivalent to 1 human genome/month)

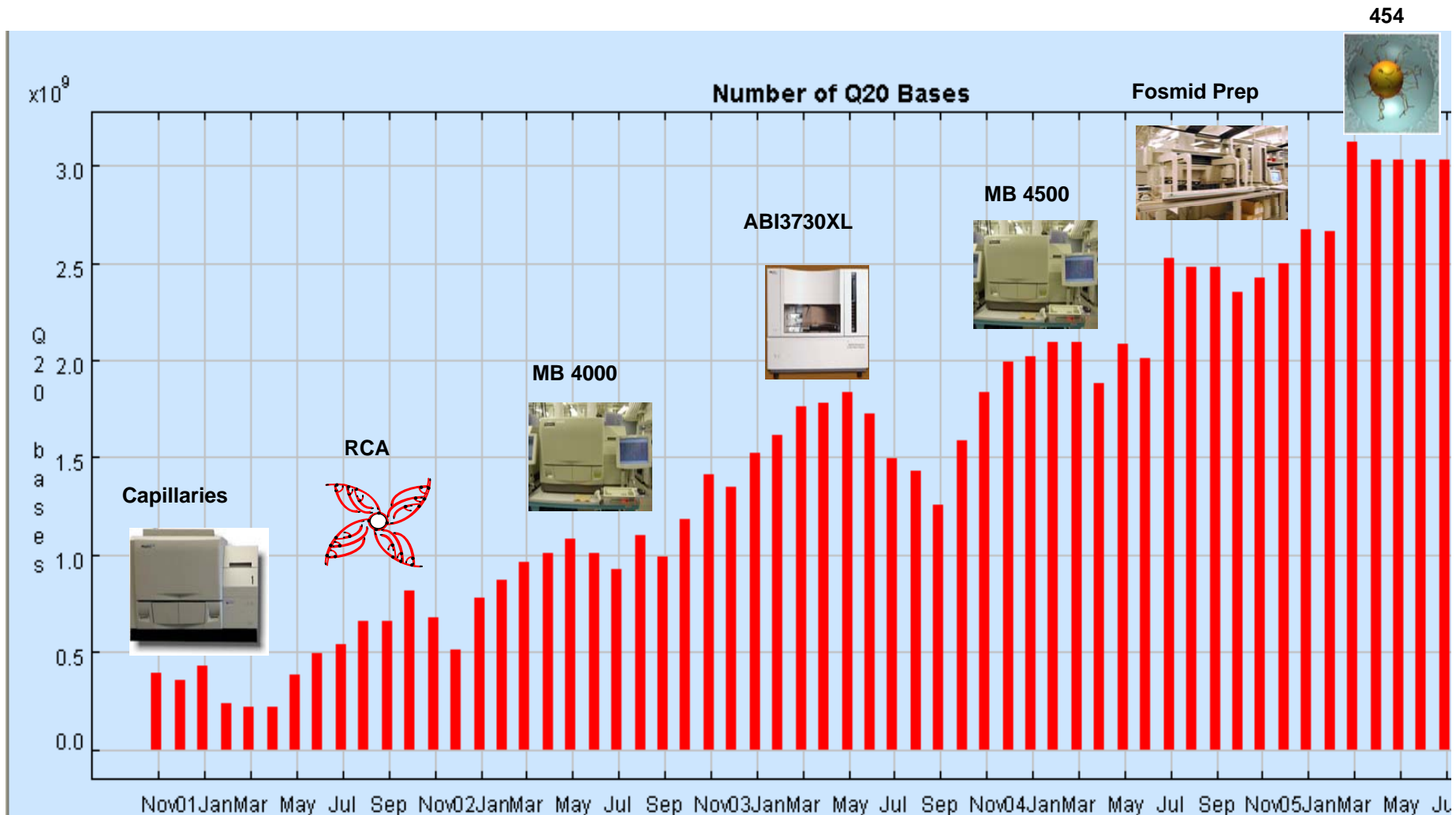
**Total (3/99-4/06) 114 Gb**

(equivalent of sequencing 38 human genomes)



**Main Production Metrics; Read Length & Pass Rate**

# New Technologies Fuel Growth in Capacity and Reduce Costs



# Capillary Sequencing

## MegaBace

1000



Qty 10; 2000 to  
Qty 84; 2002

4000



Qty 21; 2002 to Qty 36, 2003

4500



Qty 36, upgraded 2004

7 days a week operation

400 plates/day

## ABI

377



Qty 28; 1997



Qty 5; 2001  
Qty 35; 2002  
Qty 55; 2003

3730s



Qty 70; 2004

# 7yr Throughput Increases Correlated with # Sequencers

ABI

28, ABI377s

5, ABI3730s

35, ABI3730s

55, ABI3730s

70, ABI3730s

MB

20, MB1000s

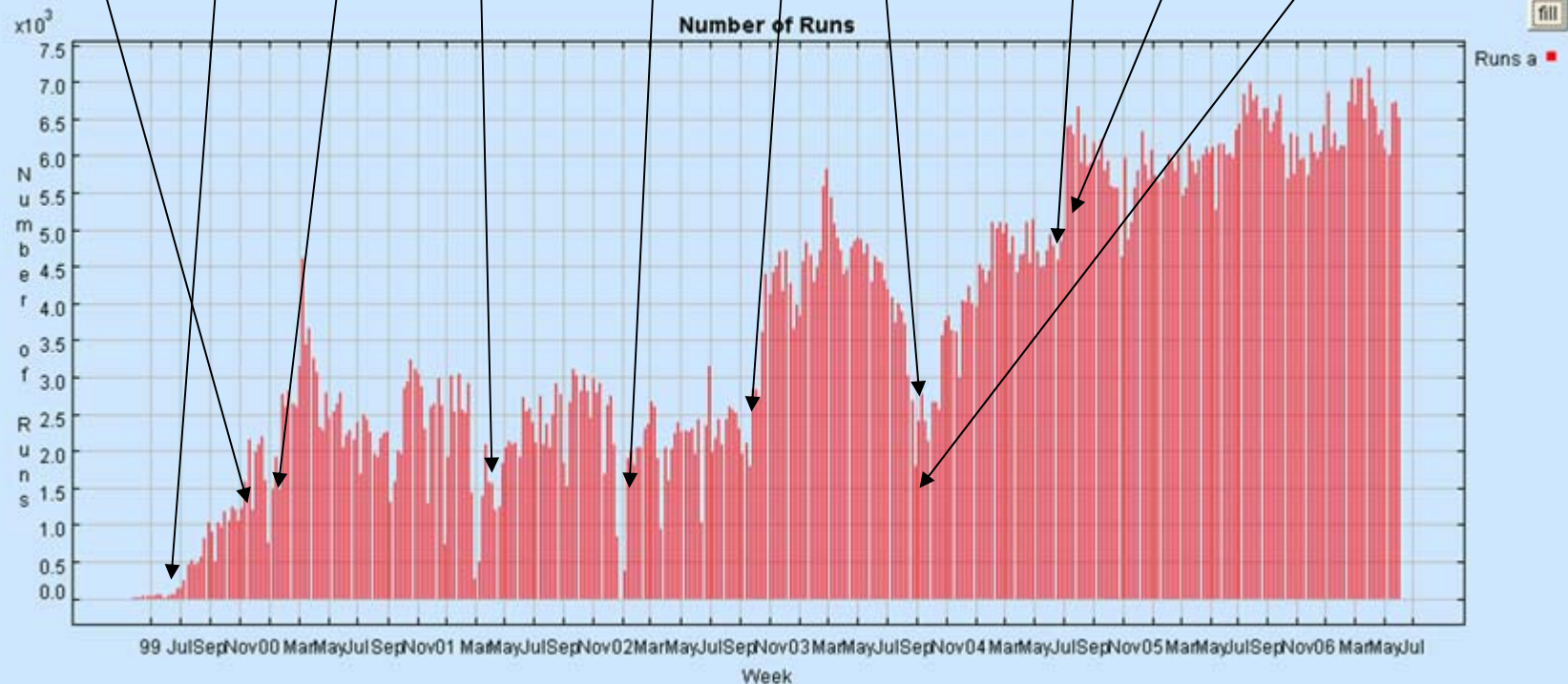
84, MB1000s

21, MB4000s

36, MB4000s

35, MB4500s

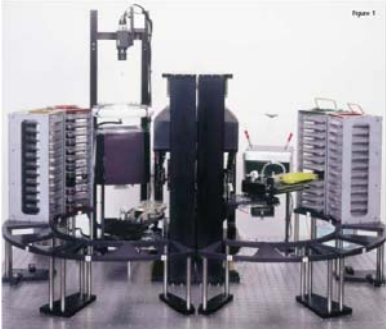
## Weekly Number of Runs Summary



TOTAL, Sequencer Type

# Colony Picking

Agar Colonies on BioAssay Plates into 384 Destination Plates



1999 Genomic Solutions & Hybaid Colony Pickers



QPix2 XT

(Qty 1; 2004)

2<sup>nd</sup> Due 7/06

**NEW**

QPix2 (Qty 1; 2000, Qty 3; 2001, Qty4; 2002)

**OLD**

Throughput ~Trays/day

96 pin picking head

Throughput (2 shifts = 16hrs)

~115 Bioassay Trays to produce 300  
(384 well) destination plates/day based  
on ~1000 colonies/tray

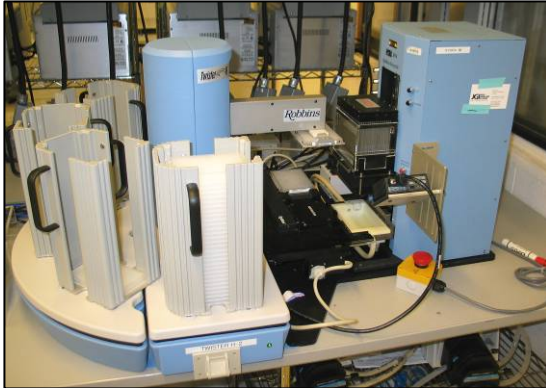


2 BioAssays & Stacked Dest Plates

# Sequence Chemistry

Aliquot RCA Product to Daughter Plates  
& Dispense FWD & REV Chemistry

**Hydra-Twister  
System**  
Qty 2; 2001



21 source to 42  
destination ~1hr/batch

**OLD**



**Cavro Syringe  
Pump Dispenser  
System**



42 plates  
~40min/batch

Typical Aliquot 1.5uL RCA Product into  
two destination plates followed by 3.5uL  
Sequencing Chemistry Cocktail Dispense

Throughput (2 shifts = 16hrs)

6 batches, 42 source plates = 84  
destination plates

500+ plates/day

**NEW**

**CyBi-Well Vario  
Integrated System**  
Qty 2; 2006



2.5hr/batch



# Why Automate?

- **In high throughput environment;**
  - automation does not necessarily increase productivity of workers,
  - does increase repeatability
  - increase equipment reliability
  - frees up operator to perform other tasks
  - reduce risk to operator of ergonomic issues
- **“Islands of Automation”**
  - modular instruments with stacks of micro-titer plates transported & loaded by operators
- **Semi-Automated Approach**
  - minimizes cost and maximizes flexibility
  - compared to fully robotic systems
  - presents a unique set of challenging issues when moving large volumes of plates



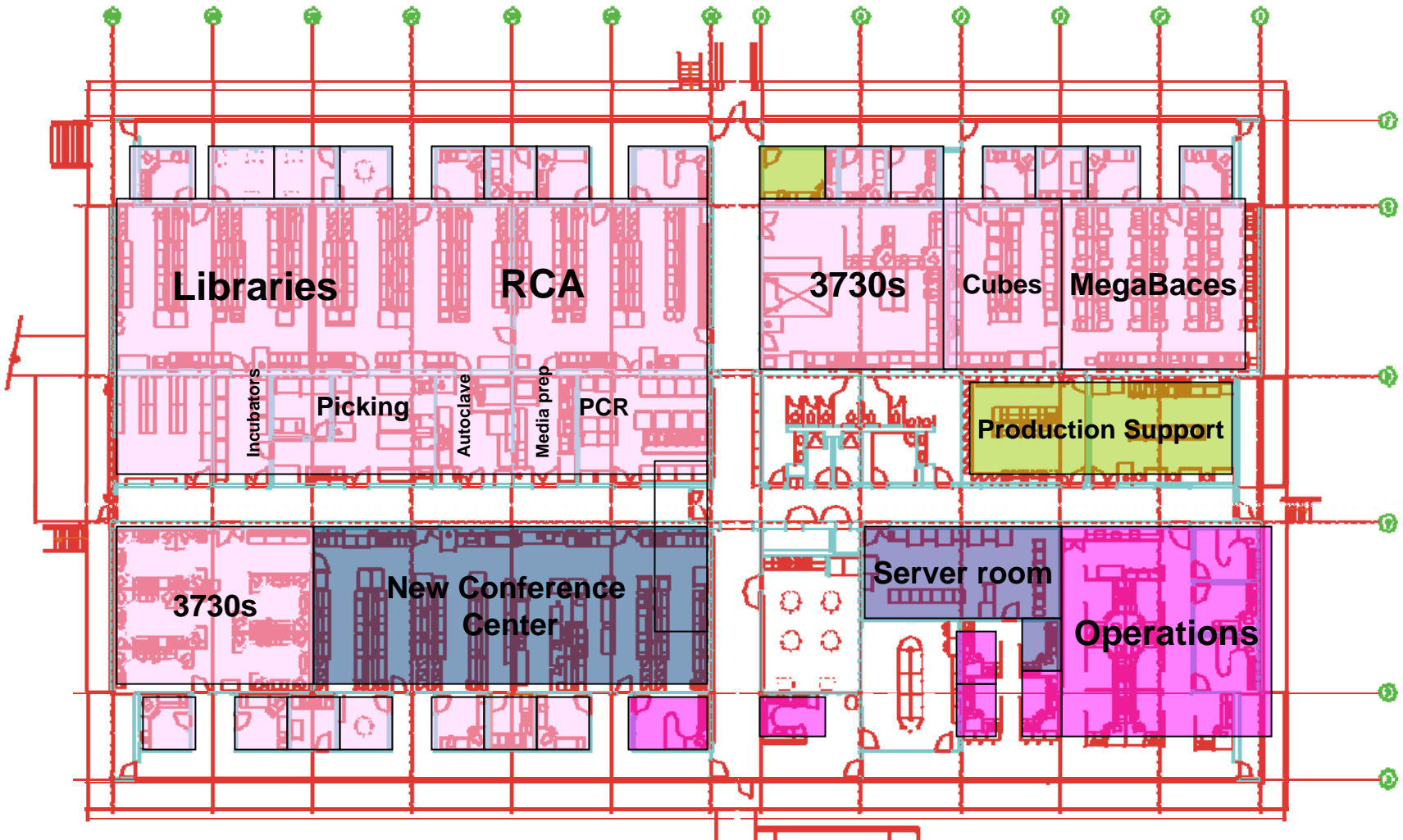
# Why Automate?

- **Need to automate driven by;**
  - Lowering Reagent Volumes of expensive reagents <1uL & increments of 0.1uL
  - Dilutions only go so far
  - Wet & dry dispense
- **1536 not going to work in the near term**
  - sequencers in 384 format
  - no thermal cycling
- **Physical constraints of laboratory space, big issue currently (next slide)**

## 2006 Community Sequencing Program



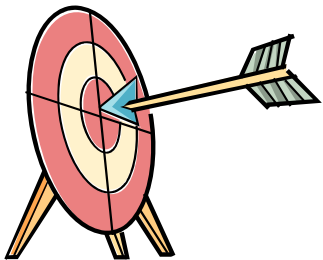
# JGI Facility Layout



# Approach to Automation Selection

- **Performance**

- Configuration, plate types (source & destination, conditioned (thermal cycled & incubated)), reagent types, throughput, software compatibility, GUI ease of use, barcode scanning, ancillary services
- Error Rate (defining major & minor), how going to test stacker, pipettor, dispenser, barcode reader, tip wash.
- Measurement device, volumes, wet or dry dispense
- Safety – seismic restraint, ergonomics, hazards, “E” stop
- Integrated system test – method outline
- Operational testing
- Precision %CV, accuracy, reliability, ergonomic design



# Approach to Automation Selection

- **Cost**
  - Expense warranted, efficient use of tax payer money
- **Delivery**
  - time, how, installation, training
- **Service**
  - field support, location of service engineers, response time, loaner parts, consumables supply



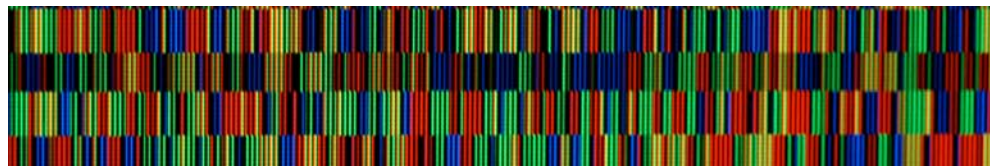
# Approach to Automation Implementation

- **Technology Transfer**
  - New or upgrading instrument technology
  - Internal process improvements
  - New processes introduced to production from R&D
- **Performance**
  - Acceptance criteria
  - Operational testing
    - Duplicate partial batch
    - Single production batch
    - Statistical Analysis of Results
- **Training**
  - SOP written during development phase, working draft
  - Operator needs to know what to expect

# Summary

- Brief history of the Joint Genome Institute
- DNA sequencing production line
- Increased throughput since 1999
- Instrumentation improvement highlights
- Our approach to successful selection & implementation

AGTCCGCGAATACAGGCTCGGT





# More Information?

**JGI**  
DOE JOINT GENOME INSTITUTE  
US DEPARTMENT OF ENERGY  
OFFICE OF SCIENCE

[search jgi](#)  
[contact us](#)  
[site map](#)  
[internal](#)

ABOUT US  
PROGRAMS  
SEQUENCING  
JGI SCIENCE  
MEETINGS  
NEWS  
EDUCATION  
EMPLOYMENT

JGI brings the expertise of four national laboratories, [Lawrence Berkeley](#), [Lawrence Livermore](#), [Los Alamos](#), and [Oak Ridge](#), and the [Stanford Human Genome Center](#) to bear on the frontiers of genome sequencing and related biology. Our sequencing targets encompass a rapidly expanding range of microbes, animals, and plants. The new [Community Sequencing Program \(CSP\)](#) aims to broaden the range still further. JGI is operated by the [University of California](#) for the [U.S. Department of Energy](#).

**genomes**

[Microbial](#), [Eukaryotic](#)  
[Integrated Microbial Genomes \(IMG\)](#)  
[IMG w/ Microbiome Samples \(IMG/M\)](#)

**latest news**

[JGI Finishes 100th Microbial Genome](#)  
[DOE BER call for bioenergy sequencing targets](#)

**sequencing**

This fiscal year : 23.407 billion base pairs sequenced  
[More statistics](#)  
[Sequencing for researchers](#)



Page last updated 5/23/2006 · [Disclaimer](#) · [Comments/Questions](#)  
 ©2006 The Regents of the University of California · [Credits](#)

More Information [www.jgi.doe.gov](http://www.jgi.doe.gov)

# Acknowledgements

- **Martin Pollard, JGI Instrumentation Manager**
- **Susan Lucas, JGI Production Manager**
- **David Gilbert, JGI Public Relations Manager**



- This work was performed under the auspices of the US Department of Energy's Office of Science, Biological and Environmental Research Program, and by the University of California, Lawrence Livermore National Laboratory under Contract No. W-7405-Eng-48, Lawrence Berkeley National Laboratory under contract No. DE-AC02-05CH11231 and Los Alamos National Laboratory under contract No. DE-AC02-06NA25396